

PATENT APPLICATION

MUSIC INTERFACE FOR MEDIA-RICH INTERACTIVE PROGRAM GUIDE

Inventor(s):

Donald F. Gordon, a citizen of United States, residing at,
465 Gabilan Street #10
Los Altos, CA 94022

Sadik Bayrakeri, a citizen of Turkey, residing at,
733 Shell Boulevard #104
Foster City, CA 94404

Arleen F. Delos Reyes, a citizen of United States, residing at,
1183 Foster City Blvd. #2
Foster City, CA 94404

Robert T. Cala, a citizen of United States residing at,
4112 Marsten Avenue
Belmont, CA 94002

Assignee:

DIVA Systems Corporation
800 Saginaw Drive
Redwood City, CA 94063

Entity: Small business concern

TOWNSEND and TOWNSEND and CREW LLP
Two Embarcadero Center, 8th Floor
San Francisco, California 94111-3834
Tel: 650-326-2400

MUSIC INTERFACE FOR MEDIA-RICH INTERACTIVE PROGRAM GUIDE

CROSS-REFERENCES TO RELATED APPLICATIONS

5 This application claims the benefit of U.S. provisional Application Serial No. 60/219,076, entitled "MUSIC INTERFACE FOR MEDIA-RICH INTERACTIVE PROGRAM GUIDE," filed July 18, 2000, which is incorporated herein by reference in its entirety for all purposes.

10 This application is further a continuation-in-part of the following U.S. Patent Applications, all which are hereby incorporated by reference in their entirety for all purposes:

Serial No. 09/583,388, entitled "ENCODING OPTIMIZATION TECHNIQUES FOR ENCODING PROGRAM GRID SECTION OF SERVER-CENTRIC IPG," filed May 30, 2000; and

15 Serial No. 09/585,263, entitled "CHANNEL INFORMATION WINDOW VIA SERVER-CENTRIC INTERACTIVE USER INTERFACE," filed June 2, 2000, which is a continuation-in-part of Serial No. 09/562,491, entitled "CHANNEL INFORMATION WINDOW VIA SERVER-CENTRIC INTERACTIVE USER INTERFACE," filed May 1, 2000, which is a continuation-in-part of Serial No.
20 09/359,560, entitled "INTERACTIVE USER INTERFACE," filed July 22, 1999.

BACKGROUND OF THE INVENTION

1. Field of the Invention

25 The present invention relates to communications systems in general and, more specifically, the invention relates to encoding techniques for use in an interactive multimedia information delivery system.

2. Description of the Background Art

30 Over the past few years, the television industry has seen a transformation in a variety of techniques by which its programming is distributed to consumers. Cable television systems are doubling or even tripling system bandwidth with the migration to hybrid fiber coax (HFC) cable transmission systems. Customers unwilling to subscribe to local cable systems have switched in high numbers to direct broadcast satellite (DBS)

systems. And, a variety of other approaches have been attempted focusing primarily on high bandwidth digital technologies, intelligent two way set top boxes, or other methods of attempting to offer service differentiated from standard cable and over the air broadcast systems.

5 With this increase in bandwidth, the number of programming choices has also increased. Leveraging off the availability of more intelligent set top boxes, several companies have developed elaborate systems for providing interactive listings. These interactive listings may include the following aspects and features: a vast array of channel offerings; expanded textual information about individual programs; the ability to
10 look forward to plan television viewing as much as several weeks in advance; and the option of automatically programming a video cassette recorder (VCR) to record a future broadcast of a television program.

 Unfortunately, the existing program guides have several drawbacks. They tend to require a significant amount of memory, some of them needing upwards of one
15 megabyte of memory at the set top terminal (STT). They are very slow to acquire their current database of programming information when they are turned on for the first time or are subsequently restarted (e.g., a large database may be downloaded to a STT using only a vertical blanking interval (VBI) data insertion technique). Disadvantageously, such slow database acquisition may result in out-of-date database information or, in the case of
20 a pay-per-view (PPV) or video-on-demand (VOD) system, limited scheduling flexibility for the information provider.

SUMMARY OF THE INVENTION

 The invention provides techniques for a music interface in a media-rich
25 interactive program guide. The music interface can be implemented using a server-centric system, and can further be implemented as an extension to the user interface for programming guide for video channels.

 An embodiment of the invention covers a method for providing an interactive music interface. In accordance with the method, an interactive program guide
30 (IPG) page having included therein a music icon representative of the music topic is initially provided. A selection for the music icon is later received and, in response, a music interface page having included therein a listing of music contents (or channels) is provided for display.

Thereafter, an indication is received that a particular music channel has been selected. In response, an audio stream associated with the selected music channel can be retrieved and processed. A music channel can be highlighted in response to movement of a cursor over the music channel and automatically selected and tuned to.

5 Alternatively, a music channel can be selected by a user via depression of a particular key or keys (e.g., a Select key, or the numeric keys for the music channel number) on a remote control unit. A data stream associated with the selected music channel may be decoded to retrieve descriptive information for the selected music channel, which can then be provided for display. A video stream associated with the selected music channel
10 may be retrieved, decoded, and provided for display.

Another embodiment of the invention provides an interactive music interface page that includes one or more (e.g., two) display regions and a channel description region. The display regions are configurable to display a listing of a group of music channels, and the channel description region is configurable to display information
15 for a particular music channel in the listing. Each display region can include a particular (e.g., alternating color strips) background and can further be configurable to display a video in place of the music channel listing. The music interface page may further include a header region configurable to display one or more objects (icons) used to identify the music interface page.

20 In constructing the music interface page, a set top terminal (STT) may retrieve (pre-generated) graphics for the page. The background and commonly used icons may be pre-loaded to the terminals. In case of icon changes, the pre-loaded icons may be periodically or occasionally sent from a head end to the STT. During construction of the music interface page, the STT may overlay pre-generated (bitmap) text within the stripes
25 of the background to create the listing of available music contents. In addition, icons may be overlayed at specific locations in the music interface page.

The invention further provides other methods and music interface pages, and STT and system (e.g., head end) that implement the methods and music interface pages described herein.

30 The foregoing, together with other aspects of this invention, will become more apparent when referring to the following specification, claims, and accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The teachings of the present invention can be readily understood by considering the following detailed description in conjunction with the accompanying drawings, in which:

5 FIG. 1 depicts an example of one frame of an interactive program guide (IPG) taken from a video sequence that may be encoded using an embodiment the present invention;

 FIG. 2 depicts a block diagram of an illustrative interactive information distribution system that may include the encoding unit and process of an embodiment of
10 the present invention;

 FIG. 3 depicts a slice map for the IPG of FIG. 1;

 FIG. 4 depicts a block diagram of the encoding unit of FIG. 2;

 FIG. 5 depicts a block diagram of the local neighborhood network of FIG.
15 2;

 FIG. 6 depicts a matrix representation of program guide data with the data groupings shown for efficient encoding;

 FIG. 7 is a diagrammatic flow diagram of a process for generating a portion of transport stream containing intra-coded video and graphics slices;

 FIG. 8 is a diagrammatic flow diagram of a process for generating a
20 portion of transport stream containing predictive-coded video and graphics slices;

 FIG. 9 illustrates a data structure of a transport stream used to transmit the IPG of FIG. 1;

 FIG. 10 is a diagrammatic flow diagram of a alternative process for generating a portion of transport stream containing predictive-coded video and graphics
25 slices;

 FIG. 11A depicts an illustration of an IPG having a graphics portion and a plurality of video portions;

 FIG. 11B depicts a slice map for the IPG of FIG. 11A;

 FIG. 12 is a diagrammatic flow diagram of a process for generating a
30 portion of transport stream containing intra-coded video and graphics slices for an IPG having a graphics portion and a plurality of video portions;

FIG. 13 is a diagrammatic flow diagram of a process for generating a portion of transport stream containing predictive-coded video and graphics slices for an IPG having a graphics portion and a plurality of video portions;

FIG. 14 depicts a block diagram of a receiver within subscriber equipment
5 suitable for use in an interactive information distribution system;

FIG. 15 depicts a flow diagram of a first embodiment of a slice recombination process;

FIG. 16 depicts a flow diagram of a second embodiment of a slice recombination process;

FIG. 17 depicts a flow diagram of a third embodiment of a slice recombination process;

FIG. 18 depicts a flow diagram of a fourth embodiment of a slice recombination process;

FIG. 19 is a schematic diagram illustrating slice-based formation of an intra-coded portion of a stream of packets including multiple intra-coded guide pages and multiple intra-coded video signals;

FIG. 20 is a schematic diagram illustrating slice-based formation of a video portion of predictive-coded stream of packets including multiple predictive-coded video signals;

FIG. 21 is a schematic diagram illustrating slice-based formation of a guide portion of predictive-coded stream of packets including skipped guide pages;

FIG. 22 is a block diagram illustrating a system and apparatus for multiplexing various packet streams to generate a transport stream;

FIG. 23 is a schematic diagram illustrating slice-based partitioning of multiple objects;

FIG. 24 is a block diagram illustrating a cascade compositor for resizing and combining multiple video inputs to create a single video output that may be encoded into a video object stream;

FIG. 25 is a block diagram illustrating a system and apparatus for multiplexing video object and audio streams to generate a transport stream;

FIG. 26 is a block diagram illustrating a system and apparatus for demultiplexing a transport stream to regenerate video object and audio streams for subsequent decoding;

FIG. 27 is a schematic diagram illustrating interacting with objects by selecting them to activate a program guide, an electronic commerce window, a video on-demand window, or an advertisement video;

FIG. 28 is a schematic diagram illustrating interacting with an object by selecting it to activate a full-resolution broadcast channel;

FIG. 29 is a flow chart illustrating an object selection operation;

FIG. 30 is a schematic diagram illustrating PID filtering prior to slice recombination;

FIG. 31 is a schematic diagram illustrating slice recombination;

FIG. 32 is a block diagram illustrating a general head-end centric system to encode and deliver a combined real time and non-real time multimedia content;

FIG. 33 depicts, in outline form, a layout 3300 of an IPG frame in accordance with an embodiment of the present invention;

FIG. 34 depicts the program grid section 3302 of the layout 3300 of fig. 33 in accordance with an embodiment of the invention;

FIG. 35 depicts an encoding process 3500 that includes low-pass filtering in accordance with an embodiment of the invention;

FIG. 36 depicts a design of an interactive program guide (IPG) page in accordance with an embodiment of the invention; and

FIGS. 37 and 38 depict two designs of a music interface page in accordance with an embodiment of the invention.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

Embodiments of the present invention relate to a system for generating, distributing and receiving a transport stream containing compressed video and graphics information. Embodiments of the present invention may be illustratively used to encode a plurality of interactive program guides (IPGs) that enable a user to interactively review, preview and select programming for a television system.

Embodiments of the present invention utilize compression techniques to reduce the amount of data to be transmitted and increase the speed of transmitting program guide information. As such, the data to be transmitted is compressed so that the available transmission bandwidth is used more efficiently. To transmit an IPG having both graphics and video, embodiments of the present invention separately encode the

graphics from the video such that the encoder associated with each portion of the IPG can be optimized to best encode the associated portion. Embodiments of the present invention may illustratively use a slice-based, predictive encoding process that is based upon the Moving Pictures Experts Group (MPEG) standard known as MPEG-2. MPEG-2 is specified in the ISO/IEC standards 13818, which is incorporated herein by reference.

The above-referenced standard describes data processing and manipulation techniques that are well suited to the compression and delivery of video, audio and other information using fixed or variable rate digital communications systems. In particular, the above-referenced standard, and other "MPEG-like" standards and techniques, compress, illustratively, video information using intra-frame coding techniques (such as run-length coding, Huffman coding and the like) and inter-frame coding techniques (such as forward and backward predictive coding, motion compensation and the like). Specifically, in the case of video processing systems, MPEG and MPEG-like video processing systems are characterized by prediction-based compression encoding of video frames with or without intra- and/or inter-frame motion compensation encoding.

To enhance error recovery, the MPEG-2 standard contemplates the use of a "slice layer" where a video frame is divided into one or more slices. A slice contains one or more contiguous sequence of macroblocks. The sequence begins and ends at any macroblock boundary within the frame. An MPEG-2 decoder, when provided a corrupted bitstream, uses the slice layer to avoid reproducing a completely corrupted frame. For example, if a corrupted bitstream is decoded and the decoder determines that the present slice is corrupted, the decoder skips to the next slice and begins decoding. As such, only a portion of the reproduced picture is corrupted.

Embodiments of the present invention may use the slice layer for the main purpose of flexible encoding and compression efficiency in a head end centric end-to-end system. A slice-based encoding system enables the graphics and video of an IPG to be efficiently coded and flexibly transmitted as described below. Consequently, a user can easily and rapidly move from one IPG page to another IPG page.

A. An Exemplary Interactive Program Guide

Embodiments of the present invention may be employed for compressing and transmitting various types of video frame sequences that contain graphics and video information, and may be particularly useful in compressing and transmitting interactive program guides (IPG) where a portion of the IPG contains video (referred to herein as the

video portion or multimedia section) and a portion of the IPG contains a programming guide grid (referred to herein as the guide portion or graphics portion or program grid section). The present invention slice-based encodes the guide portion separately from the slice-based encoded video portion, transmits the encoded portions within a transport stream, and reassembles the encoded portions to present a subscriber (or user) with a comprehensive IPG. Through the IPG, the subscriber can identify available programming and select various services provided by their information service provider.

FIG. 1 depicts a frame from an illustrative IPG page 100. In this particular embodiment of an IPG, the guide grid information is contained in portion 102 (left half page) and the video information is contained in portion 101 (right half page). The IPG display 100 comprises: first 105A, second 105B and third 105C time slot objects; a plurality of channel content objects 110-1 through 110-8; a pair of channel indicator icons 141A, 141B; a video barker 120 (and associated audio barker); a cable system or provider logo 115; a program description region 150; a day of the week identification object 131; a time of day object 139; a next time slot icon 134; a temporal increment/decrement object 132; a "favorites" filter object 135, a "movies" filter object 136; a "kids" (i.e., juvenile) programming filter icon 137; a "sports" programming filter object 138; and a VOD programming icon 133. It should be noted that the day of the week object 131 and next time slot icon 134 may comprise independent objects (as depicted in FIG. 1) or may be considered together as parts of a combined object.

A user may transition from one IPG page to another, where each page contains a different graphics portion 102, i.e., a different program guide graphics. The details regarding the encoding and decoding of a series of IPG pages in accordance with the present invention are provided below.

Details regarding the operation of the IPG page of FIG. 1, the interaction of this page with other pages and with a user are described in commonly assigned US patent application no. 09/359,560 filed July 23, 1999 which is hereby incorporated herein by reference.

B. System

FIG. 2 depicts a high-level block diagram of an information distribution system 200, e.g., a video-on-demand system or digital cable system, which may incorporate an embodiment of the present invention. The system 200 contains head end equipment (HEE) 202, local neighborhood equipment (LNE) 228, a distribution network

204 (e.g., hybrid fiber-coax network) and subscriber equipment (SE) 206. This form of information distribution system is disclosed in commonly assigned U.S. patent application serial number 08/984,710 filed December 3, 1997. The system is known as DIVA™ provided by DIVA Systems Corporation.

5 The HEE 202 produces a plurality of digital streams that contain encoded information in illustratively MPEG-2 compressed format. These streams are modulated using a modulation technique that is compatible with a communications channel 230 that couples the HEE 202 to one or more LNE (in FIG. 1, only one LNE 228 is depicted). The LNE 228 is illustratively geographically distant from the HEE 202. The LNE 228 selects
10 data for subscribers in the LNE's neighborhood and remodulates the selected data in a format that is compatible with distribution network 204. Although the system 200 is depicted as having the HEE 202 and LNE 228 as separate components, those skilled in the art will realize that the functions of the LNE may be easily incorporated into the HEE202. It is also important to note that the presented slice-based encoding method is
15 not constrained to physical location of any of the components. The subscriber equipment (SE) 206, at each subscriber location 2061, 2062, , 206n, comprises a receiver 224 and a display 226. Upon receiving a stream, the subscriber equipment receiver 224 extracts the information from the received signal and decodes the stream to produce the information on the display, i.e., produce a television program, IPG page, or other multimedia program.

20 In an interactive information distribution system such as the one described in commonly assigned U.S. patent application 08/984,710, filed December 3, 1997, the program streams are addressed to particular subscriber equipment locations that requested the information through an interactive menu. A related interactive menu structure for requesting video-on-demand is disclosed in commonly assigned U.S. patent application
25 serial number 08/984,427, filed December 3, 1997. Another example of interactive menu for requesting multimedia services is the interactive program guide (IPG) disclosed in commonly assigned U.S. patent application 60/093,891, filed in July 23, 1998.

 To assist a subscriber (or other viewer) in selecting programming, the HEE
202 produces information that can be assembled to create an IPG such as that shown in
30 FIG. 1. The HEE produces the components of the IPG as bitstreams that are compressed for transmission in accordance with the present invention.

 A video source 214 supplies the video sequence for the video portion of the IPG to an encoding unit 216 of the present invention. Audio signals associated with the video sequence are supplied by an audio source 212 to the encoding and multiplexing

unit 216. Additionally, a guide data source 232 provides program guide data to the encoding unit 216. This data is typically in a database format, where each entry describes a particular program by its title, presentation time, presentation date, descriptive information, channel, and program source.

5 The encoding unit 216 compresses a given video sequence into one or more elementary streams and the graphics produced from the guide data into one or more elementary streams. As described below with respect to FIG. 4, the elementary streams are produced using a slice-based encoding technique. The separate streams are coupled to the cable modem 222.

10 The streams are assembled into a transport stream that is then modulated by the cable modem 222 using a modulation format that is compatible with the head end communications channel 230. For example, the head end communications channel may be a fiber optic channel that carries high-speed data from the HEE 202 to a plurality of LNE 228. The LNE 228 selects IPG page components that are applicable to its
15 neighborhood and re-modulates the selected data into a format that is compatible with a neighborhood distribution network 204. A detailed description of the LNE 228 is presented below with respect to FIG. 5.

 The subscriber equipment 206 contains a receiver 224 and a display 226 (e.g., a television). The receiver 224 demodulates the signals carried by the distribution
20 network 204 and decodes the demodulated signals to extract the IPG pages from the stream. The details of the receiver 224 are described below with respect to FIG. 14.

C. Encoding Unit 216

 The system of the present invention is designed specifically to work in a
25 slice-based ensemble encoding environment, where a number of bitstreams are generated to compress video information using a sliced-based technique. In the MPEG-2 standard, a "slice layer" may be created that divides a video frame into one or more "slices". Each slice includes one or more macroblocks, where the macroblocks are illustratively defined as rectangular groups of pixels that tile the entire frame, e.g., a frame may consist of 30
30 rows and 22 columns of macroblocks. Any slice may start at any macroblock location in a frame and extend from left to right and top to bottom through the frame. The stop point of a slice can be chosen to be any macroblock start or end boundary. The slice layer syntax and its conventional use in forming an MPEG-2 bitstream is well known to those skilled in the art and shall not be described herein.

When the invention is used to encode an IPG comprising a graphics portion and a video portion, the slice-based technique separately encodes the video portion of the IPG and the grid graphics portion of the IPG. As such, the grid graphics portion and the video portion are represented by one or more different slices. FIG. 3 illustrates an exemplary slice division of an IPG 100 where the guide portion 102 and the video portion 101 are each divided into N slices (e.g., g/s1 through g/sN and v/s1 through v/sN). Each slice contains a plurality of macroblocks, e.g., 22 macroblocks total and 11 macroblocks in each portion. The slices in the graphics portion are pre-encoded to form a "slice form grid page" database that contains a plurality of encoded slices of the graphics portion. The encoding process can also be performed real-time during the broadcast process depending on the preferred system implementation. In this way, the graphics slices can be recalled from the database and flexibly combined with the separately encoded video slices to transmit the IPG to the LNE and, ultimately, to the subscribers. The LNE assembles the IPG data for the neighborhood as described below with respect to FIG. 5.

Although the following description is presented within the context of an IPG, it is important to note that the present invention may be equally applicable in a broad range of applications, such as: broadcast video on demand delivery; e-commerce; Internet video education services; and similar applications.

As depicted in FIG. 4, the encoding unit 216 receives a video sequence and an audio signal. The audio source comprises, illustratively, audio information that is associated with a video portion in the video sequence such as an audio track associated with still or moving images. For example, in the case of a video sequence representing a movie trailer, the audio stream is derived from the source audio (e.g., music and voice-over) associated with the movie trailer.

The encoding unit 216 comprises video processor 400, a graphics processor 402 and a controller 404. The video processor 400 comprises a compositor unit 406 and an encoder unit 408. The compositor unit 406 combines a video sequence with advertising video, advertiser or service provider logos, still graphics, animation, or other video information. The encoder unit 408 comprises one or more video encoders 410, e.g., a real-time MPEG-2 encoder and an audio encoder 412, e.g., an AC-3 encoder. The encoder unit 408 produces one or more elementary streams containing slice-based encoded video and audio information.

The video sequence is coupled to a real time video encoder 410. The video encoder then forms a slice-based bitstream, e.g., an MPEG-2 compliant bit stream, for the video portion of an IPG. For purposes of this discussion, it is assumed that the GOP structure consists of an I-picture followed by ten B-pictures, where a P-picture separates each group of two B-pictures (i.e., "I-B-B-P-B-B-P-B-B-P-B-B-P-B-B"), however, any GOP structure and size may be used in different configurations and applications.

The video encoder 410 "pads" the graphics portion (illustratively the left half portion of IPG) with null data. The null data may be replaced by the graphics grid slices, at a later step, within the LNE. Since the video encoder processes only motion video information, excluding the graphics data, it is optimized for motion video encoding.

The controller 404 manages the slice-based encoding process such that the video encoding process is time and spatially synchronized with the grid encoding process. This is achieved by defining slice start and stop locations according to the objects in the IPG page layout and managing the encoding process as defined by the slices.

The graphics portion of the IPG is separately encoded in the graphics processor 402. The processor 402 is supplied guide data from the guide data source (232 in FIG. 2). Illustratively, the guide data is in a conventional database format containing program title, presentation date, presentation time, program descriptive information and the like. The guide data grid generator 414 formats the guide data into a "grid", e.g., having a vertical axis of program sources and a horizontal axis of time increments. One specific embodiment of the guide grid is depicted and discussed in detail above with respect to FIG. 1.

The guide grid is a video frame that is encoded using a video encoder 416 optimized for video with text and graphics content. The video encoder 416, which can be implemented as software, slice-based encodes the guide data grid to produce one or more bitstreams that collectively represent the entire guide data grid. The encoder is optimized to effectively encode the graphics and text content.

The controller 404 defines the start and stop macroblock locations for each slice. The result is a GOP structure having intra-coded pictures containing I-picture slices and predicted pictures containing B and P-picture slices. The I-pictures slices are separated from the predicted picture slices. Each encoded slice is separately stored in a slice form grid page database 418. The individual slices can be addressed and recalled

from the database 418 as required for transmission. The controller 404 controls the slice-based encoding process as well as manages the database 418.

D. Local Neighborhood Equipment (LNE) 228

5 FIG. 5 depicts a block diagram of the LNE 228. The LNE 228 comprises a cable modem 500, slice combiner 502, a multiplexer 504 and a digital video modulator 506. The LNE 228 is coupled illustratively via the cable modem to the HEE 202 and receives a transport stream containing the encoded video information and the encoded guide data grid information. The cable modem 500 demodulates the signal from the HEE
10 202 and extracts the MPEG slice information from the received signal. The slice combiner 502 combines the received video slices with the guide data slices in the order in which the decoder at receiver side can easily decode without further slice re-organization. The resultant combined slices are PID assigned and formed into an illustratively MPEG compliant transport stream(s) by multiplexer 504. The slice-combiner (scanner) and
15 multiplexer operation is discussed in detail with respect to FIGS. 5-10. The transport stream is transmitted via a digital video modulator 506 to the distribution network 204.

 The LNE 228 is programmed to extract particular information from the signal transmitted by the HEE 202. As such, the LNE can extract video and guide data grid slices that are targeted to the subscribers that are connected to the particular LNE.
20 For example, the LNE 228 can extract specific channels for representation in the guide grid that are available to the subscribers connected to that particular LNE. As such, unavailable channels to a particular neighborhood would not be depicted in a subscriber's IPG. Additionally, the IPG can contain targeted advertising, e-commerce, program notes, and the like. As such, each LNE can combine different guide data slices with different
25 video to produce IPG screens that are prepared specifically for the subscribers connected to that particular LNE. Other LNEs would select different IPG component information that is relevant to their associated subscribers.

 FIG. 6 illustrates a matrix representation 600 of a series of IPG pages. In the illustrated example, ten different IPG pages are available at any one time period, e.g.,
30 t1, t2, and so on. Each page is represented by a guide portion (g) and a common video portion (v) such that a first IPG page is represented by g1/v1, the second IPG page is represented by g2/v1 and so on. In the illustrative matrix 600, ten identical guide portions (g1-g10) are associated with a first video portion (v1). Each portion is slice-base encoded as described above within the encoding unit (216 of FIG.4).

FIG. 6 illustrates the assignment of PIDs to the various portions of the IPG pages. In the figure, only the content that is assigned a PID is delivered to a receiver. The intra-coded guide portion slices g1 through g10 are assigned to PID1 through PID10 respectively. One of the common intra-coded video portion v1, illustratively the tenth IPG page, is assigned to PID11. In this form, substantial bandwidth saving is achieved by delivering intra-coded video portion slices v1 only one time. Lastly, the predictive-coded slices g1/v2 through g1/v15 are assigned to PID11. As shown in the figure, a substantial bandwidth saving is achieved by transmitting only one group of illustratively fourteen predicted picture slices, g1/v2 to g1/v15. This is provided by the fact that the prediction error images for each IPG page 1 to 10 through time units t2 to t15 contain the same residual images. Further details of PID assignment process are discussed in next sections.

FIG. 7 depicts a process 700 that is used to form a bitstream 710 containing all the intra-coded slices encoded at a particular time t1 of FIG. 6. At step 702, a plurality of IPG pages 7021 through 70210 are provided to the encoding unit. At step 704, each page is slice base encoded to form, for example, guide portion slices g1/s1 through g1/sN and video portion slices v/s1 through v/sN for IPG page 1 7041. The slice based encoding process for video and guide portions can be performed in different forms. For example, guide portion slices can be pre-encoded by a software MPEG-2 encoder or encoded by the same encoder as utilized for encoding the video portion. If the same encoder is employed, the parameters of the encoding process are adjusted dynamically for both portions. It is important to note that regardless of the encoder selection and parameter adjustment, each portion is encoded independently. While encoding the video portion, the encoding is performed by assuming the full frame size (covering both guide and video portions) and the guide portion of the full frame is padded with null data. This step, step 704, is performed at the HEE. At step 706, the encoded video and guide portion slices are sent to the LNE. If the LNE functionality is implemented as part of the HEE, then, the slices are delivered to the LNE as packetized elementary stream format or any similar format as output of the video encoders. If LNE is implemented as a remote network equipment, the encoded slices are formatted in a form to be delivered over a network via a preferred method such as cable modem protocol or any other preferred method. Once the slice-based streams are available in the LNE, the slice combiner at step 706 orders the slices in a form suitable for the decoding method at the receiver equipment. As depicted in FIG. 7 (b), the guide portion and video portion slices are ordered in a manner as if the original pictures in FIG. 7 (a) are scanned from left to right

and top to bottom order. Each of the slice packets are then assigned PID's as discussed in FIG. 6 by the multiplexer; PID1 is assigned to g1/s1 ... g1/sn, PID2 to g2/s1 ... g2/sn, ..., PID10 to g10/s1 ... g10/sn, and PID11 is assigned to v/s1 ... v/sn. The resultant transport stream containing the intra-coded slices of video and guide portions is illustrated in FIG.

- 5 7 (c). Note that based on this transport stream structure, a receiving terminal as discussed in later parts of this description of the invention, retrieves the original picture by constructing the video frames row-by-row, first retrieving, assuming PID1 is desired, e.g., g1/s1 of PID1 then v/s1 of PID11, next g1/s2 of PID1 then v/s2 of PID11 and so on.

- FIG. 8 illustrates a process 800 for producing a bitstream 808 containing
- 10 the slices from the predictive-coded pictures accompanying the transport stream generation process discussed in FIG. 7 for intra-coded slices. As shown in FIG. 6, illustratively, only the predicted slices belonging to IPG page 1 is delivered. Following the same arguments of encoding process in FIG. 7, at step 802, the predictive-coded slices are generated at the HEE independently and then forwarded to an LNE either as local or
- 15 in a remote network location. At step 804, slices in the predictive-coded guide and video portion slices, illustratively from time periods t2 to t15, are scanned from left to right and top to bottom in slice-combiner and complete data is assigned PID 11 by the multiplexer. Note that the guide portion slices g1/s1 to g1/sn at each time period t2 to t15 does not change from their intra-coded corresponding values at t1. Therefore, these slices are
- 20 coded as skipped macroblocks "sK". Conventional encoder systems do not necessarily skip macroblocks in a region even when there is no change from picture to picture. At step 806, the slice packets are ordered into a portion of final transport stream, first including the video slice packets v2/s1 ... v2/SN to v15/s1 ... v15/sN, then including the skipped guide slices sK/s1 ... sK/sN from t2 to t15 in the final transport stream. FIG. 9
- 25 depicts a complete MPEG compliant transport stream 900 that contains the complete information needed by a decoder to recreate IPG pages that are encoded in accordance with the invention. The transport stream 900 comprises the intra-coded bitstream 710 of the guide and video slices (PIDS1 to 11), a plurality of audio packets 902 identified by an audio PID, and the bitstream 806 containing the predictive-coded slices in PID11. The
- 30 rate of audio packet insertion between video packets is decided based on the audio and video sampling ratios. For example, if audio is digitally sampled as one tenth of video signal, then an audio packet may be introduced into the transport stream every ten video packets. The transport stream 900 may also contain, illustratively after every 64 packets, data packets that carry to the set top terminal overlay updates, raw data, HTML, java,

URL, instructions to load other applications, user interaction routines, and the like. The data PIDs are assigned to different set of data packets related to guide portion slice sets and also video portion slice sets.

FIG. 10 illustrates a process 1000, an alternative embodiment of process 800 depicted in FIG. 8, for producing a predictive-coded slice bitstream 1006. The process 1000, at step 1002, produces the slice base encoded predictive-coded slices. At step 1004, the slices are scanned to intersperse the "skipped" slices (sk) with the video slices (v1). The previous embodiment scanned the skipped guide portion and video portion separately. In this embodiment, each slice is scanned left to right and top to bottom completely, including the skipped guide and video data. As such, at step 1008, the bitstream 1006 has the skipped guide and video slices distributed uniformly throughout the transport stream.

The foregoing embodiments assumed that the IPG page was divided into one guide portion and one video portion. For example, in FIG. 1, the guide portion is the left half of the IPG page and the video portion is the right half of the IPG page. However, the invention can be extended to have a guide portion and multiple video portions, e.g., three. Each of the video portions may contain video having different rates of motion, e.g., portion one may run at 30 frames per second, portions two and three may run at 2 frames per second. FIG. 11A illustrates an exemplary embodiment of an IPG 1100 having a guide portion 1102 and three video portions 1104, 1106 and 1108. To encode such an IPG, each portion is separately encoded and assigned PIDs. FIG. 11B illustrates an assignment map for encoding each portion of the IPG page of FIG. 11A. The guide portion 1002 is encoded as slices g/s1 through g/sN, while the first video portion 1004 is encoded as slices v/s1 through v/sM, and the second video portion 1006 is encoded as slices j/sM+1 through j/sL, the third video portion 1008 is encoded as slices p/sL+1 through p/sN.

FIG. 12 depicts the scanning process 1200 used to produce a bitstream 1210 containing the intra-coded slices. The scanning process 1200 flows from left to right, top to bottom through the assigned slices of FIG. 11B. PIDs are assigned, at step 1202, to slices 1 to M; at step 1204, to slices M+1 to L; and, at step 1206, to slices L+1 to N. As the encoded IPG is scanned, the PIDS are assigned to each of the slices. The guide portion slices are assigned PIDS 1 through 10, while the first video portion slices are assigned PID11, the second video portion slices are assigned PID12 and the third video portion slices are assigned PID13. The resulting video portion of the bitstream 1210

contains the PIDS for slices 1-M, followed by PIDS for slices M+1 to L, and lastly by the PIDS for L+1 to N.

FIG. 13 depicts a diagrammatical illustration of a process 1300 for assigning PIDS to the predictive-coded slices for the IPG of FIG. 11A. The scanning process 1300 is performed, at step 1302, from left to right, top to bottom through the V, J and P predicted encoded slices and PIDS are assigned where the V slices are assigned PID11, the J slices are assigned PID 12 and the P slices are assigned PID13. After the video portion predicted encoded slices have assigned PIDs, the process 1300, at step 1304, assigns PIDs to the skipped slices. The skipped guide slices vertically corresponding to the V slices are assigned PID11, the skipped slices vertically corresponding to the J slices are assigned PID12 and the skipped slices vertically corresponding to the P slices are assigned PID13. At step 1308, the resulting predictive-coded bitstream 1312 comprises the predicted video slices in portion 1306 and the skipped slices 1310. The bitstream 1210 of intra-coded slices and the bitstream 1312 of predictive-coded slices are combined into a transport stream having a form similar to that depicted in FIG. 9.

To change pages in the guide, it is required to switch between programs (video PIDs for groups of slices) in a seamless manner. This cannot be done cleanly using a standard channel change by the receiver switching from PID to PID directly, because such an operation flushes the video and audio buffers and typically gives half a second blank screen.

To have seamless decoder switching, a splice countdown (or random access indicator) method is employed at the end of each video sequence to indicate the point at which the video should be switched from one PID to another.

Using the same profile and constant bit rate coding for the video and graphics encoding units, the generated streams for different IPG pages are formed in a similar length compared to each other. This is due to the fact that the source material is almost identical differing only in the characters in the guide from one page to another. In this way, while streams are generated having nearly identical lengths, the streams are not exactly the same length. For example, for any given sequence of 15 video frames, the number of transport packets in the sequence varies from one guide page to another. Thus, a finer adjustment is required to synchronize the beginnings and ends of each sequence across all guide pages in order for the countdown switching to work.

Synchronization of a plurality of streams may be accomplished in a way that provides seamless switching at the receiver.

Three methods are provided for that purpose:

First, for each sequence the multiplexer in the LNE identifies the length of the longest guide page for that particular sequence, and then adds sufficient null packets to the end of each other guide page so that all the guide pages become the same length. Then, the multiplexer adds the switching packets at the end of the sequence, after all the null packets.

The second method requires buffering of all the packets for all guide pages for each sequence. If this is allowed in the considered system, then the packets can be ordered in the transport stream such that the packets for each guide page appear at slightly higher or lower frequencies, so that they all finish at the same point. Then, the switching packets are added by the multiplexer in the LNE at the end of each stream without the null padding.

A third method is to start each sequence together, and then wait until all the packets for all the guide pages have been generated. Once the generation of all packets is completed, switching packets are placed in the streams at the same time and point in each stream.

Depending on the implementation of decoder units within the receiver and requirements of the considered application, each one of the methods can be applied with advantages. For example, the first method, which is null-padding, can be applied to avoid bursts of N packets of the same PID into a decoder's video buffer faster than the MPEG specified rate (e.g., 1.5 Mbit).

The teachings of the above three methods can be extended apply to similar synchronization problems and to derive similar methods for ensuring synchronization during stream switching.

E. Receiver 224

FIG. 14 depicts a block diagram of the receiver 224 (also known as a set top terminal (STT) or user terminal) suitable for use in producing a display of an IPG in accordance with the present invention. The STT 224 comprises a tuner 1410, a demodulator 1420, a transport demultiplexer 1430, an audio decoder 1440, a video decoder 1450, an on-screen display processor (OSD) 1460, a frame store memory 1462, a video compositor 1490 and a controller 1470. User interaction is provided via a remote

control unit 1480. Tuner 1410 receives, e.g., a radio frequency (RF) signal comprising, for example, a plurality of quadrature amplitude modulated (QAM) information signals from a downstream (forward) channel. Tuner 1410, in response to a control signal TUNE, tunes a particular one of the QAM information signals to produce an intermediate frequency (IF) information signal. Demodulator 1420 receives and demodulates the intermediate frequency QAM information signal to produce an information stream, illustratively an MPEG transport stream. The MPEG transport stream is coupled to a transport stream demultiplexer 1430.

Transport stream demultiplexer 1430, in response to a control signal TD produced by controller 1470, demultiplexes (i.e., extracts) an audio information stream A and a video information stream V. The audio information stream A is coupled to audio decoder 1440, which decodes the audio information stream and presents the decoded audio information stream to an audio processor (not shown) for subsequent presentation. The video stream V is coupled to the video decoder 1450, which decodes the compressed video stream V to produce an uncompressed video stream VD that is coupled to the video compositor 1490. OSD 1460, in response to a control signal OSD produced by controller 1470, produces a graphical overlay signal VOSD that is coupled to the video compositor 1490. During transitions between streams representing the user interfaces, buffers in the decoder are not reset. As such, the user interfaces seamlessly transition from one screen to another.

The video compositor 1490 merges the graphical overlay signal VOSD and the uncompressed video stream VD to produce a modified video stream (i.e., the underlying video images with the graphical overlay) that is coupled to the frame store unit 1462. The frame store unit 562 stores the modified video stream on a frame-by-frame basis according to the frame rate of the video stream. Frame store unit 562 provides the stored video frames to a video processor (not shown) for subsequent processing and presentation on a display device.

Controller 1470 comprises a microprocessor 1472, an input/output module 1474, a memory 1476, an infrared (IR) receiver 1475 and support circuitry 1478. The microprocessor 1472 cooperates with conventional support circuitry 1478 such as power supplies, clock circuits, cache memory and the like as well as circuits that assist in executing the software routines that are stored in memory 1476. The controller 1470 also contains input/output circuitry 1474 that forms an interface between the controller 1470 and the tuner 1410, the transport demultiplexer 1430, the onscreen display unit 1460, the

back channel modulator 1495, and the remote control unit 1480. Although the controller 1470 is depicted as a general-purpose computer that is programmed to perform specific interactive program guide control function in accordance with the present invention, the invention can be implemented in hardware as an application specific integrated circuit (ASIC). As such, the process steps described herein are intended to be broadly interpreted as being equivalently performed by software, hardware, or a combination thereof.

In the exemplary embodiment of FIG. 14, the remote control unit 1480 comprises an 8-position joystick, a numeric pad, a "select" key, a "freeze" key and a "return" key. User manipulations of the joystick or keys of the remote control device are transmitted to a controller via an infrared (IR) link. The controller 1470 is responsive to such user manipulations and executes related user interaction routines 1400, uses particular overlays that are available in an overlay storage 1479.

After the signal is tuned and demodulated, the video streams are recombined via stream processing routine 1402 to form the video sequences that were originally compressed. The processing unit 1402 employs a variety of methods to recombine the slice-based streams, including, using PID filter 1404, demultiplexer 1430, as discussed in the next sections of this disclosure of the invention. Note that the PID filter implemented illustratively as part of the demodulator is utilized to filter the undesired PIDs and retrieve the desired PIDs from the transport stream. The packets to be extracted and decoded to form a particular IPG are identified by a PID mapping table (PMT) 1477. After the stream processing unit 1402 has processed the streams into the correct order (assuming the correct order was not produced in the LNE), the slices are sent to the MPEG decoder 1450 to generate the original uncompressed IPG pages. If an exemplary transport stream with two PIDs as discussed in previous parts of the this disclosure, excluding data and audio streams, is received, then the purpose of the stream processing unit 1402 is to recombine the intra-coded slices with their corresponding predictive-coded slices in the correct order before the recombined streams are coupled to the video decoder. This complete process is implemented as software or hardware. In the illustrated IPG page slice structure, only one slice is assigned per row and each row is divided into two portions, therefore, each slice is divided into guide portion and video portion. In order for the receiving terminal to reconstruct the original video frames, one method is to construct a first row from its two slices in the correct order by retrieving two corresponding slices from the transport stream, then construct a second row from its two

slices, and so on. For this purpose, a receiver is required to process two PIDs in a time period. The PID filter can be programmed to pass two desired PIDs and filter out the undesired PIDs. The desired PIDs are identified by the controller 1470 after the user selects an IPG page to review. A PID mapping table (1477 of FIG. 14) is accessed by the controller 1470 to identify which PIDs are associated with the desired IPG. If a PID filter is available in the receiver terminal, then it is utilized to receive two PIDs containing slices for guide and video portions. The demultiplexer then extracts packets from these two PIDs and couples the packets to the video decoder in the order in which they arrived. If the receiver does not have an optional PID filter, then the demultiplexer performs the two PID filtering and extracting functions. Depending on the preferred receiver implementation, the following methods are provided in FIGS. 15-18 to recombine and decode slice-based streams.

E1. Recombination Method 1

In this first method, intra-coded slice-based streams (I-streams) and the predictive-coded slice-based streams (PRED streams) to be recombined keep their separate PID's until the point where they must be depacketized. The recombination process is conducted within the demultiplexer 1430 of the subscriber equipment. For illustrative purposes, assuming a multi-program transport stream with each program consisting of I-PIDs for each intra-coded guide slice, I-PIDs for the intra-coded video slices, one PRED-PID for predicted guide and video, an audio-PID, and multiple data-PIDs, any packet with a PID that matches any of the PID's within the desired program (as identified in a program mapping table) are depacketized and the payload is sent to the elementary stream video decoder. Payloads are sent to the decoder in exactly in the order in which the packets arrive at the demultiplexer.

FIG. 15 is a flow diagram of the first packet extraction method 1500. The method starts at step 1505 and proceeds to step 1510 to wait for (user) selection of an I-PID to be received. The I-PID, as the first picture of a stream's GOP, represents the stream to be received. However, since the slice-based encoding technique assigns two or more I-PIDS to the stream (i.e., I-PIDs for the guide portion and for one or more video portions), the method must identify two or more I-PIDs. Upon detecting a transport packet having the selected I-PIDs, the method 1500 proceeds to step 1515.

At step 1515, the I-PID packets (e.g., packets having PID-1 and PID-11) are extracted from the transport stream, including the header information and data, until

the next picture start code. The header information within the first-received I-PID access unit includes sequence header, sequence extension, group start code, GOP header, picture header, and picture extension, which are known to a reader that is skilled in MPEG-1 and MPEG-2 compression standards. The header information in the next I-PID access units
5 that belongs to the second and later GOP's includes group start code, picture start code, picture header, and extension. The method 1500 then proceeds to step 1520 where the payloads of the packets that includes header information related to video stream and I-picture data are coupled to the video decoder 1550 as video information stream V. The method 1500 then proceeds to step 1525.

10 At step 1525, the predicted picture slice-based stream packets PRED-PID, illustratively the PID-11 packets of fourteen predicted pictures in a GOP of size fifteen, are extracted from the transport stream. At step 1530, the payloads of the packets that include header information related to video stream and predicted-picture data are coupled to the video decoder 1550 as video information stream V. At the end of step 1530, a
15 complete GOP, including the I-picture and the predicted-picture slices, are available to the video decoder 1550. As the payloads are sent to the decoder in exactly in the order in which the packets arrive at the demultiplexer, the video decoder decodes the recombined stream with no additional recombination process. The method 1500 then proceeds to step 1535.

20 At step 1535, a query is made as to whether a different I-PID is requested, e.g., new IPG is selected. If the query at step 1535 is answered negatively, then the method 1500 proceeds to step 1510 where the transport demultiplexer 1530 waits for the next packets having the PID of the desired I-picture slices. If the query at step 1535 is answered affirmatively, then the PID of the new desired I-picture slices is identified at
25 step 1540 and the method 1500 returns to step 1510.

The method 1500 of FIG. 15 is used to produce a conformant MPEG video stream V by concatenating a desired I-picture slices and a plurality of P- and/or B-picture slices forming a pre-defined GOP structure.

30 E2. Recombination Method 2

The second method of recombining the video stream involves the modification of the transport stream using a PID filter. A PID filter 1404 can be implemented as part of the demodulator 1420 of FIG. 14 or as part of demultiplexer.

For illustrative purposes, assuming a multi-program transport stream with each program consisting of an I-PIDs for both video and guide, PRED-PID for both video and guide, audio-PID, and data-PID, any packet with a PID that matches any of the PIDs within the desired program as identified by the program mapping table to be received have its PID modified to the lowest video PID in the program (the PID which is referenced first in the program's program mapping table (PMT)). For example, in a program, assuming that a guide slice I-PID is 50, the video slice I-PID is 51 and PRED-PID is 52. Then, the PID-filter modifies the video I-PID and the PRED-PID as 50 and thereby, I- and Predicted-Picture slice access units attain the same PID number and become a portion of a common stream.

As a result, the transport stream output from the PID filter contains a program with a single video stream, whose packets appear in the proper order to be decoded as valid MPEG bitstream.

Note that the incoming bit stream does not necessarily contain any packets with a PID equal to the lowest video PID referenced in the programs PMT. Also note that it is possible to modify the video PID's to other PID numbers than lowest PID without changing the operation of the algorithm.

When the PID's of incoming packets are modified to match the PID's of other packets in the transport stream, the continuity counters of the merged PID's may become invalid at the merge points, due to each PID having its own continuity counter. For this reason, the discontinuity indicator in the adaptation field is set for any packets that may immediately follow a merge point. Any decoder components that check the continuity counter for continuity is required to correctly process the discontinuity indicator bit.

FIG. 16 illustrates the details of this method, in which, it starts at step 1605 and proceeds to step 1610 to wait for (user) selection of two I-PIDs, illustratively two PIDs corresponding to guide and video portion slices, to be received. The I-PIDs, comprising the first picture of a stream's GOP, represents the two streams to be received. Upon detecting a transport packet having one of the selected I-PIDs, the method 1600 proceeds to step 1615.

At step 1615, the PID number of the I-stream is re-mapped to a predetermined number, PID*. At this step, the PID filter modifies all the PID's of the desired I-stream packets to PID*. The method then proceeds to step 1620, wherein the PID number of the predicted picture slice streams, PRED-PID, is re-mapped to PID*. At

this step, the PID filter modifies all the PID's of the PRED-PID packets to PID*. The method 1600 then proceeds to step 1625.

At step 1625, the packets of the PID* stream are extracted from the transport stream by the demultiplexer. The method 1600 then proceeds to step 1630, where the payloads of the packets that includes video stream header information and I-picture and predicted picture slices are coupled to the video decoder as video information stream V. Note that the slice packets are ordered in the transport stream in the same order as they are to be decoded, i.e., a guide slice packets of first row followed by video slice packets of first row, second row, and so on. The method 1600 then proceeds to 1635.

At step 1635, a query is made as to whether a different set of (two) I-PIDs is requested. If the query at step 1635 is answered negatively, then the method 1600 proceeds to step 1610 where the transport demultiplexer waits for the next packets having the identified I-PIDs. If the query at step 1635 is answered affirmatively, then the two PIDs of the new desired I-picture is identified at step 1640 and the method 1600 returns to step 1610.

The method 1600 of FIG. 16 is used to produce a conformant MPEG video stream by merging the intra-coded slice streams and predictive-coded slice streams before the demultiplexing process.

E3. Recombination Method 3

The third method accomplishes MPEG bitstream recombination by using splicing information in the adaptation field of the transport packet headers by switching between video PIDs based on splice countdown concept.

In this method, the MPEG streams signal the PID to PID switch points using the splice countdown field in the transport packet header's adaptation field. When the PID filter is programmed to receive one of the PIDs in a program's PMT, the reception of a packet containing a splice countdown value of 0 in its header's adaptation field causes immediate reprogramming of the PID filter to receive the other video PID. Note that a special attention to splicing syntax is required in systems where splicing is used also for other purposes.

FIG. 17 illustrates the details of this method, in which, it starts at step 1705 and proceeds to step 1710 to wait for (user) selection of two I-PIDs to be received. The I-PIDs, comprising the first picture of a stream's GOP, represents the stream to be received.

Upon detecting a transport packet having one of the selected I-PIDs, the method 1700 proceeds to step 1715.

At step 1715, the I-PID packets are extracted from the transport stream until, and including, the I-PID packet with slice countdown value of zero. The method 1700 then proceeds to step 1720 where the payloads of the packets that includes header information related to video stream and I-picture slice data are coupled to the video decoder as video information stream V. The method 1700 then proceeds to step 1725.

At step 1725, the PID filter is re-programmed to receive the predicted picture packets PRED-PID. The method 1700 then proceeds to 1730. At step 1730, the predicted stream packets, illustratively the PID11 packets of predicted picture slices, are extracted from the transport stream. At step 1735, the payloads of the packets that include header information related to video stream and predicted-picture data are coupled to the video decoder. At the end of step 1735, a complete GOP, including the I-picture slices and the predicted-picture slices, are available to the video decoder. As the payloads are sent to the decoder in exactly in the order in which the packets arrive at the demultiplexer, the video decoder decodes the recombined stream with no additional recombination process. The method 1700 then proceeds to step 1740.

At step 1740, a query is made as to whether a different I-PID set (two) is requested. If the query at step 1740 is answered negatively, then the method 1700 proceeds to step 1750 where the PID filter is re-programmed to receive the previous desired I-PIDs. If answered affirmatively, then the PIDs of the new desired I-picture is identified at step 1745 and the method proceeds to step 1750, where the PID filter is re-programmed to receive the new desired I-PIDs. The method then proceeds to step 1745, where the transport demultiplexer waits for the next packets having the PIDs of the desired I-picture.

The method 1700 of FIG. 17 is used to produce a conformant MPEG video stream, where the PID to PID switch is performed based on a splice countdown concept. Note that the slice recombination can also be performed by using the second method where the demultiplexer handles the receiving PIDs and extraction of the packets from the transport stream based on the splice countdown concept. In this case, the same process is applied as FIG. 17 with the difference that instead of reprogramming the PID filter after "0" splice countdown packet, the demultiplexer is programmed to depacketize the desired PIDs.

E4. Recombination Method 4

For the receiving systems that do not include a PID filter and for those receiving systems in which the demultiplexer cannot process two PIDs for splicing the streams, a fourth method presented herein provides the stream recombination. In a receiver that cannot process two PIDs, two or more streams with different PIDs are spliced together via an additional splicing software or hardware and can be implemented as part of the demultiplexer. The process is described below with respect to FIG. 18. The algorithm provides the information to the demultiplexer about which PID to be spliced to as the next step. The demultiplexer processes only one PID but a different PID after the splice occurs.

FIG. 18 depicts a flow diagram of this fourth process 1800 for recombining the IPG streams. The process 1800 begins at step 1801 and proceeds to step 1802 wherein the process defines an array of elements having a size that is equal to the number of expected PIDs to be spliced. It is possible to distribute splice information in a picture as desired according to slice structure of the picture and the desired processing form at the receiver. For example, in the slice based streams discussed in this invention, for an I picture, splice information may be inserted into slice row portions of guide and video data. At step 1804, the process initializes the video PID hardware with for each entry in the array. At step 1810, the hardware splice process is enabled and the packets are extracted by the demultiplexer. The packet extraction may also be performed at another step within the demultiplexer. At step 1812, the process checks a hardware register to determine if a splice has been completed. If the splice has occurred, the process, at step 1814, disables the splice hardware and, at step 1816, sets the video PID hardware to the next entry in the array. The process then returns along path 1818 to step 1810. If the splice has not occurred, the process proceeds to step 1820 wherein the process waits for a period of time and then returns along path 1822 to step 1812.

In this manner, the slices are spliced together by the hardware within the receiver. To facilitate recombining the slices, the receiver is sent an array of valid PID values for recombining the slices through a user data in the transport stream or another communications link to the STT from the HEE. The array is updated dynamically to ensure that the correct portions of the IPG are presented to the user correctly. Since the splice points in slice based streams may occur at a frequent level, a software application may not have the capability to control the hardware for splicing operation as discussed

above. If this is the case, then, firmware is dedicated to control the demodulator hardware for splicing process at a higher rate than a software application can handle.

F. Example: Interactive Program Guide

5 The video streams representing the IPG may be carried in a single transport stream or multiple transport streams, within the form of a single or multi-programs as discussed below with respect to the description of the encoding system. A user desiring to view the next 1.5 hour time interval (e.g., 9:30 – 11:00) may activate a “scroll right” object (or move the joystick to the right when a program within program
10 grid occupies the final displayed time interval). Such activation results in the controller of the STT noting that a new time interval is desired. The video stream corresponding to the new time interval is then decoded and displayed. If the corresponding video stream is within the same transport stream (i.e., a new PID), then the stream is immediately decoded and presented. If the corresponding video stream is within a different transport
15 stream, then the related transport stream is extracted from the broadcast stream and the related video stream is decoded and presented. If the corresponding transport stream is within a different broadcast stream, then the related broadcast stream is tuned, the corresponding transport stream is extracted, and the desired video stream is decoded and presented.

20 Note that each extracted video stream is associated with a common audio stream. Thus, the video/audio barker function of the program guide is continuously provided, regardless of the selected video stream. Also note that the teachings of the invention are equally applicable to systems and user interfaces that employs multiple audio streams.

25 Similarly, a user interaction resulting in a prior time interval or a different set of channels results in the retrieval and presentation of a related video stream. If the related video stream is not part of the broadcast video streams, then a pointcast session is initiated. For this purpose, the STT sends a request to the head end via the back channel requesting a particular stream. The head end then processes the request, retrieves the
30 related guide and video streams from the information server, incorporates the streams within a transport stream as discussed above (preferably, the transport stream currently being tuned/selected by the STT) and informs the STT which PIDs should be received, and from which transport stream should be demultiplexed. The STT then extracts the related PIDs for the IPG. In the case of the PID being within a different transport stream,

the STT first demultiplexes the corresponding transport stream (possibly tuning a different QAM stream within the forward channel).

Upon completion of the viewing of the desired stream, the STT indicates to the head end that it no longer needs the stream, whereupon the head end tears down the pointcast session. The viewer is then returned to the broadcast stream from which the pointcast session was launched.

Note that the method and apparatus described herein is applicable to any number of slice assignments to a video frame and any type of slice structures. The presented algorithms are also applicable to any number of PID assignments to intra-coded and predictive-coded slice based streams. For example, multiple PIDs can be assigned to the predictive-coded slices without loss of generality. Also note that the method and apparatus described herein is fully applicable picture based encoding by assigning each picture only to a one slice, where each picture is encoded then as a full frame instead of multiple slices.

G. Multi-Functional User Interface with Picture-in-Picture Functionality

Picture-in-picture (PIP) functionality may be provided using slice-based encoding. The PIP functionality supplies multiple (instead of singular) video content. Moreover, an additional user interface (UI) layer may be provided on top (presented to the viewer as an initial screen) of the interactive program guide (IPG). The additional UI layer extends the functionality of the IPG from a programming guide to a multi-functional user interface. The multi-functional user interface may be used to provide portal functionality to such applications as electronic commerce, advertisement, video-on-demand, and other applications.

A matrix representation of IPG data with single video content is described above in relation to Fig. 6. As shown in Fig. 6, single video content, including time-sequenced video frames V1 to V15, is shared among multiple guide pages g1 to g10. A diagrammatic flow of a slice-based process for generating a portion of the transport stream containing intra-coded video and graphics slices is described above in relation to Fig. 7. As described below, slice-based encoding may also be used to provide picture-in-picture (PIP) functionality and a multi-functional user interface.

FIG. 19 is a schematic diagram illustrating slice-based formation of an intra-coded portion of a stream of packets 1900 including multiple intra-coded guide pages and multiple intra-coded video frames. The intra-coded video frames generally

occur at a first frame of a group of pictures (GOP). Hence, the schematic diagram in Fig. 19 is denoted as corresponding to time t_1 .

In the example illustrated in Fig. 19, packet identifiers (PIDs) 1 through 10 are assigned to ten program guide pages (g1 through g10), and PIDs 11 through 13 are assigned to three video streams (V1, M1, and K1). Each guide page is divided into N slices S1 to SN, each slice extending from left to right of a row. Likewise, each intra-coded video frame is divided into N slices s1 to sN.

As shown in Fig. 19, one way to form a stream of packets is to scan guide and video portion slices serially. In other words, packets from the first slice (s1) are included first, then packets from the second slice (s2) are included second, then packets from the third slice (s3) are included third, and so on until packets from the Nth slice (sN) are included last, where within each slice grouping, packets from the guide graphics are included in serial order (g1 to g10), then packets from the intra-coded video slices are included in order (V1, M1, K1). Hence, the stream of packets is included in the order illustrated in Fig. 19.

FIG. 20 is a schematic diagram illustrating slice-based formation of predictive-coded portion of multiple video stream packets. The predictive-coded video frames (either predicted P or bidirectional B frames in MPEG2) generally occur after the first frame of a group of pictures (GOP). For Fig. 20, it is assumed that the GOP has 15 frames. Hence, the schematic diagram in Fig. 20 is denoted as corresponding to times t_2 to t_{15} .

In the example illustrated in Fig. 20, PIDs 11 through 13 are assigned to three video streams (V1, M1, and K1), each predictive-coded video frame of each video stream being divided into N slices s1 to sN.

As shown in Fig. 20, one way to form a stream of packets is to scan serially from the time t_2 through t_N . In other words, packets 2002 from the second time (t_2) are included first, then packets 2003 from the third time (t_3) are included second, then packets 2004 from the fourth time (t_4) are included third, and so on until packets 2015 from the fifteenth time (t_{15}) are included last. Within each time, packets of predictive-coded video frames from each video stream are grouped together by slice (S1 through S15). Within each slice grouping, the packets are ordered with the packet corresponding to the slice for video stream V as first, the packet corresponding to the slice for video stream M as second, and the packet corresponding to the slice for video

stream K as third. Hence, the stream of packets is included in the order illustrated in Fig. 20.

FIG. 21 is a schematic diagram illustrating slice-based formation of a stream of packets including skipped guide pages. The formation of the stream of packets in Fig. 21 is similar to the formation of the stream of packets in Fig. 20. However, the skipped guide page content (SK) is the same for each slice and for each video stream. In contrast, the predictive-coded video frames are different for each slice and for each video stream.

For each time t2 through t15, the packets containing the skipped guide pages may follow the corresponding packets containing the predictive-coded video frames. For example, for time t2, the first row of skipped guide packets 2102 follow the first row of predictive-coded packets 2002. For time t3, the second row of skipped guide packets 2103 follow the second row of predictive-coded packets 2003. And so on.

FIG. 22 is a block diagram illustrating a system and apparatus for multiplexing various packet streams to generate a transport stream. The apparatus shown in Fig. 22 may be employed as part of the local neighborhood equipment (LNE) 228 of the distribution system described above in relation to Fig. 2. In the example illustrated in Fig. 22, the various packet streams include three packetized audio streams 2202, 2204, and 2206, and the video and graphic packet stream 2214 comprising the intra-coded 1900, predictive-coded 2000, and skipped-coded 2100 packets.

The three packetized audio streams 2202, 2204, and 2206 are input into a multiplexer 2208. The multiplexer 2208 combines the three streams into a single audio packet stream 2210. The single audio stream 2210 is then input into a remultiplexer 2212. An alternate embodiment of the present invention may input the three streams 2202, 2204, and 2206 directly into the remultiplexer 2212, instead of first creating the single audio stream 2210.

The video and graphic packet stream 2214 is also input into the remultiplexer 2212. As described above in relation to Figs. 19-21, the video and graphic packet stream 2214 comprises the intra-coded 1900, predictive-coded 2000, and skipped-coded 2100 packets. One way to order the packets for a single GOP is illustrated in Fig. 22. First, the packets 1900 with PID 1 to PID 13 for intra-coded guide and video at time t1 are transmitted. Second, packets 2002 with PID 11 to PID 13 for predictive-coded video at time t2 are transmitted, followed by packets 2102 with PID 11 to PID 13 for skipped-coded guide at time t2. Third, packets 2003 with PID 11 to PID 13 for predictive-

coded video at time t3 are transmitted, followed by packets 2103 with PID 11 to PID 13 for skipped-coded guide at time t3. And so on, until lastly for the GOP, packets 2015 with PID 11 to PID 13 for predictive-coded video at time t15 are transmitted, followed by packets 2115 with PID 11 to PID 13 for skipped-coded guide at time t15.

5 The remultiplexer 2212 combines the video and graphic packet stream 2214 with the audio packet stream 2210 to generate a transport stream 2216. In one embodiment, the transport stream 2216 interleaves the audio packets with video and graphics packets. In particular, the interleaving may be done such that the audio packets for time t1 are next to the video and graphics packets for time t1, the audio packets for
10 time t2 are next to the video and graphics packets for time t2, and so on.

FIG. 23 is a schematic diagram illustrating slice-based partitioning of multiple objects of an exemplary user interface that is presented to the user as an initial screen. In the example illustrated in Fig. 23, nine objects O1 through O9 are shown. As illustrated in part (a) on the left side of Fig. 23, these nine objects may be displayed on
15 one full-size video screen by dividing the screen into a 3x3 matrix with nine areas. In this case, each of the nine objects would be displayed at 1/3 of the full horizontal resolution and 1/3 of the full vertical resolution.

Part (b) on the right side of Fig. 23 shows one way for slice-based partitioning of the nine objects being displayed in the 3x3 matrix. The frame in Fig. 23(b)
20 is divided into 3N horizontal slices. Slices 1 to N include objects O1, O2, and O3, dividing each object into N horizontal slices. Slices N+1 to 2N include objects O4, O5, and O6, dividing each object into N horizontal slices. Lastly, slices 2N+1 to 3N include objects O7, O8, and O9, dividing each object into N horizontal slices.

FIG. 24 is a block diagram illustrating a cascade compositor for resizing
25 and combining multiple video inputs to create a single video output that may be encoded into a video object stream. In the example shown in Fig. 24, the number of multiple video inputs is nine. In this case, each video input corresponds to a video object from the arrangement shown in Fig. 23(a).

The first compositor 2402 receives a first set of three full-size video inputs
30 that correspond to the first row of video objects O1, O2, and O3 in Fig. 23(a). The first compositor 2402 resizes each video input by one third in each dimension, then arranges the resized video inputs to form the first row of video objects. The first compositor 2402 outputs a first composite video signal 2403 that includes the first row of video objects.

The second compositor 2404 receives the first composite video signal 2403 from the first compositor 2402. The second compositor 2404 also receives a second set of three full-size video inputs that corresponds to the second row of video objects O4, O5, and O6 in Fig. 23(a). The second compositor resizes and arranges these three video inputs. It then adds them to the first composite video signal 2403 to form a second composite video signal 2405 that includes the first and second rows of objects.

The third compositor 2406 receives the second composite video signal 2405 and a third set of three full-size video inputs that corresponds to the third row of video objects O7, O8, and O9 in Fig. 23(a). The third compositor 2406 resizes and arranges these three video inputs. It then adds them to the second composite video signal 2405 to form a third composite video signal 2407 that includes all three rows of objects.

An encoder 2408 receives the third composite video signal 2407 and digitally encodes it to form a video object stream 2409. The encoding may be slice-based encoding using the partitioning shown in Fig. 23(b).

FIG. 25 is a block diagram illustrating a system and apparatus for multiplexing video object and audio streams to generate a transport stream. The apparatus shown in Fig. 25 may be employed as part of the local neighborhood equipment (LNE) 228 of the distribution system described above in relation to Fig. 2. In the example illustrated in Fig. 25, the various packet streams include a video object stream 2502 and a multiplexed packetized audio stream 2504.

The multiplexed packetized audio stream 2504 includes multiple audio streams that are multiplexed together. Each audio stream may belong to a corresponding video object. The multiplexed packetized audio stream 2504 is input into a remultiplexer (remux) 2506.

The video object stream 2502 is also input into the remultiplexer 2506. The encoding of the video object stream 2502 may be slice-based encoding using the partitioning shown in Fig. 23(b). In this case, each object is assigned a corresponding packet identifier (PID). For example, the first object O1 is assigned PID 101, the second object O2 is assigned PID 102, the third object O3 is assigned PID 103, and so on, and the ninth object O9 is assigned PID 109.

The remultiplexer 2506 combines the video object stream 2502 with the multiplexed packetized audio stream 2504 to generate an object transport stream 2508. In one embodiment, the object transport stream 2508 interleaves the audio packets with video object packets. In particular, the interleaving may be done such that the audio

packets for time t1 are next to the video object packets for time t1, the audio packets for time t2 are next to the video object packets for time t2, and so on.

FIG. 26 is a block diagram illustrating a system and apparatus for demultiplexing a transport stream to regenerate video object and audio streams for subsequent decoding. The system and apparatus includes a demultiplexer 2602 and a video decoder 2604.

The demultiplexer 2602 receives the object transport stream 2508 and demultiplexes the stream 2508 to separate out the video object stream 2502 and the multiplexed packetized audio stream 2504. The video object stream 2502 is further processed by the video decoder 2604. For example, as illustrated in Fig. 26, the video decoder 2604 may output a video object page 2606 which displays reduced-size versions of the nine video objects O1 through O9.

FIG. 27 is a schematic diagram illustrating interaction with objects by selecting them to activate a program guide, an electronic commerce window, a video on-demand window, or an advertisement video. In the example illustrated in Fig. 27, a video display 2702 may display various objects, including multiple video channel objects (Channels A through F, for example), an advertisement object, a video on-demand (VOD) object, and an electronic commerce (e-commerce) object.

Each of the displayed objects may be selected by a user interacting with a set-top terminal. For example, if the user selects the channel A object, then the display may change to show a relevant interactive program guide (IPG) page 2704. The relevant IPG page 2704 may include, for example, a reduced-size version of the current broadcast on channel A and guide data with upcoming programming for channel A or the guide page where channel A is located. The audio may also change to the audio stream corresponding to channel A.

As another example, if the user selects the advertisement object, then the display may change to show a related advertisement video (ad video) 2706. Further, this advertisement video may be selected, leading to an electronic commerce page relating to the advertisement. The audio may also change to an audio stream corresponding to the advertisement video.

As yet another example, if the user selects the VOD object, then the display may change to show a VOD window 2708 that enables and facilitates selection of VOD content by the user. Further, once the user selects a particular video for on-demand

display, an electronic commerce page may be displayed to make the transaction between the user and the VOD provider.

As yet another example, if the user selects the electronic commerce (e-commerce) object, then the display may change to show an e-commerce window 2710 that enables and facilitates electronic commerce. For example, the e-commerce window 2710 may comprise a hypertext markup language (HTML) page including various multimedia content and hyperlinks. The hyperlinks may, for example, link to content on the world wide web, or link to additional HTML pages which provides further product information or opportunities to make transactions.

FIG. 28 is a schematic diagram illustrating interacting with an object by selecting it to activate a full-resolution broadcast channel. In this example, if the user selects the object for channel E, the display changes to a full-resolution display 2802 of the video broadcast for channel E, and the audio changes to the corresponding audio stream. The same principle applies when the channel is pointcast to a specific viewer.

FIG. 29 is an exemplary flow chart illustrating an object selection operation. While in the receiving operation, the PID filter is employed as an example to fulfill the PID selection operation, any of the preferred filtering and demultiplexing methods discussed in FIGS. 15, 16, 17, and 18 can be utilized. The exemplary operation includes the following steps:

In a first step 2902, the video decoder 2604 (decodes and) outputs the video object page 2606 that includes the nine objects O1 through O9. In a second step 2904, a user selects an object via a set top terminal or remote control. For example, the object may be the first object O1 that may correspond to channel A. In this example, selection of the first object O1 results in the display on a corresponding IPG page 2704 including guide data and a reduced-size version of the channel A broadcast.

In a third step 2906, a PID filter is reprogrammed to receive packets for O1 and associated guide data. For example, if packets for video object O1 are identified by PID 101, and packets for the associated guide data are identified by PID 1, then the PID filter would be reprogrammed to receive packets with PID 101 and PID 1. This filtering step 2906 is described further below in relation to Fig. 30. Such reprogramming of the PID filter would occur only if such a PID filter. One system and method using such a PID filter is described above in relation to Fig. 17. The methods in FIG. 15, 16, or 18 can be employed depending on the receiving terminal capabilities and requirements.

In a fourth step 2908, a demultiplexer (Demux) depacketizes slices of the first object O1 and associated guide data. Note that this step 2908 and the previous step 2906 are combined in some of the related methods of FIGS. 15, 16, and 18. Subsequently, in a fifth step 2910, a slice recombiner reconstitutes the IPG page including the reduced-size version of the channel A broadcast and the associated guide data. Slices would only be present if the first object O1 and associated guide data were encoded using a slice-based partitioning technique, such as the one described above in relation to Fig. 23(b).

Finally, in a sixth step 2912, a video decoder decodes and outputs the IPG page for viewing by the user.

FIG. 30 is a schematic diagram illustrating PID filtering prior to slice recombination. Fig. 30 shows an example of a transport stream 3002 received by a set top terminal. The transport stream 3002 includes intra-coded guide packets 3004, predictive-coded (skipped) guide packets 3006, and intra-coded and predictive-coded video object packets 3008.

In the example illustrated in Fig. 30, the intra-coded guide packets 3004 include slice-partitioned guide graphics data for the first frame of each group of pictures (GOP) for each of ten IPG pages. These intra-coded packets 3004 may, for example, be identified by PID 1 through PID 10 as described above in relation to Fig. 19.

Similarly, the skipped-coded guide packets 3006 include skipped-coded data for the second through last frames of each GOP for each of ten IPG pages. These skipped-coded packets 3006 may be identified, for example, by PID 11 as described above in relation to Fig. 21.

In the example illustrated in Fig. 30, the intra-coded and predictive-coded video object packets 3008 include slice-partitioned video data for each of nine objects O1 through O9. These packets 3008 may, for example, be identified by PID 101 through PID 109 as described above in relation to Fig. 25.

The transport stream 3002 is filtered 3010 by a PID filter. The filtering process 3010 results in received packets 3012. For example, if the PID filter is programmed to receive only packets corresponding to the first object O1 (PID 101) and associated guide data (PIDs 1 and 11), then the received packets 3012 would include only those packets with PIDs 101, 1, and 11.

FIG. 31 is a schematic diagram illustrating slice recombination. In this embodiment, slice recombination occurs after PID filtering. A slice recombiner receives the PID-filtered packets 3012 and performs the slice recombination process 3102 in

which slices are combined to form frames. As a result of the slice recombination process 3102, an intra-coded frame 3104 is formed for each GOP from the slices of the intra-coded guide page (PID 1) and the slices of the intra-coded video frame (PID 101).

Furthermore, the second to last predictive-coded frames 3106 are formed for each GOP
5 from the slices of the skipped-coded guide page (PID 11) and the slices of the predictive-coded video frames (PID 101). The above-discussed methods can be equally applied to frame-based encoding and delivery by defining a slice as a complete frame without loss of generality.

The above discussed encoding and delivery methods for PIP utilizes a
10 combination of broadcast/demandcast traffic model where multiple video signals are broadcast and delivered to the set top box even the viewer does not utilize some of the video content at a particular time. Such an approach makes response times far more consistent, and far less sensitive to the number of subscribers served. Typical latencies may remain sub-second even when the subscriber count in a single modulation group
15 (aggregation of nodes) exceeds 10 thousand. On the other hand, the bandwidth necessary to delivery the content increases compared to a point-to-point traffic model. However, with the advantage of the slice-based recombinant MPEG compression techniques, the latency reduction of broadcast/demandcast model is achieved without much bandwidth compromise.

20 In addition, with a server-centric content generation and control, the transport streams containing tremendous motion video information is delivered and decoded directly through the transport demultiplexer and MPEG decoder without being accessible to the microprocessor, saving processing and memory resources and costs at set top terminal.

25 The multi-functional user interface supports any combination of full-motion video windows, at least one or more of these video inputs can be driven from existing ad-insertion equipment enabling the operator to leverage existing equipment and infrastructure, including ad traffic and billing systems, to quickly realize added revenues. The discussed system does not have any new requirements for ad production. The ads can
30 be the same as are inserted into any other broadcast channels.

H. General Head-End Centric System Architecture for Encoding and Delivery of Combined Realtime and Non-Realtime Content

A unique feature of the head-end centric system discussed in previous sections (for encoding and delivery of interactive program guide, multi-functional user interfaces, picture-in-picture type of applications) is the combined processing of realtime and non-realtime multimedia content. In other words, the discussed head-end centric system architecture can be utilized for other related applications that contain realtime and non-realtime content in similar ways with the teachings of this invention. For further clarification, FIG. 32 illustrates a general system and apparatus for encoding, multiplexing, and delivery of realtime and non-realtime content in accordance with the present invention including: a non-realtime content source for providing non-realtime content; a non-realtime encoder for encoding the non-realtime content into encoded non-realtime content; a realtime content source for providing realtime video and audio content; a realtime encoder for encoding the realtime video and audio content into encoded realtime video and audio; a remultiplexer for repacketizing the encoded non-realtime content and the encoded realtime video and audio into transport packets; and a re-timestamp unit coupled to the remultiplexer for providing timestamps to be applied to the transport packets in order to synchronize the realtime and non-realtime content therein.

Fig. 32 is a block diagram illustrating such a system for re-timestamping and rate control of realtime and non-realtime encoded content in accordance with an embodiment of the present invention.

The apparatus includes a non-realtime content source 3202, a realtime content source, a non-realtime encoder 3206, a rate control unit 3208, a realtime encoder 3210 (including a realtime video encoder 3211 and a realtime audio encoder 3212), a slice combiner 3214, a remultiplexer 3216, a re-timestamp unit 3218, and a clock unit 3220. The apparatus shown in Fig. 32 may be included in a head-end of a cable distribution system.

The non-realtime content may include guide page graphics content for an interactive program guide (IPG). The realtime content may include video and audio advertisement content for insertion into the IPG.

The rate control unit 3208 may implement an algorithm that sets the bit rate for the output of the non-realtime encoder 3206. Based on a desired total bit rate, the algorithm may subtract out a maximum bit rate anticipated for the realtime video and

audio encoded signals. The resultant difference would basically give the allowed bit rate for the output of the non-realtime encoder 106. In a slice-based embodiment, this allowed bit rate would be divided by the number of slices to determine the allowed bit rate per slice of the IPG content. In a page-based embodiment, this allowed bit rate would be the
5 allowed bit rate per page of the IPG content.

The re-timestamp unit 3218 may receive a common clock signal from the common clock unit 3220 and generates therefrom presentation and decoding timestamps. These timestamps are transferred to the remultiplexer (Remux) 3216 for use in re-timestamping the packets (overriding existing timestamps from the encoders 3206, 3211,
10 and 3212). The re-timestamping synchronizes the non-realtime and realtime content so that non-realtime and realtime content intended to be displayed in a single frame are displayed at the same time.

The common clock unit 3220 may also provide a common clock stream to the set-top terminals. The common clock stream is transmitted in parallel with the
15 transport stream.

I. Techniques for Encoding Program Grid Section of IPG

FIG. 33 depicts, in outline form, a layout 3300 of an IPG frame in accordance with an embodiment of the present invention. The layout 3300 includes a
20 program grid section 3301 and a multimedia section 3302. The layout 300 in Fig. 33 corresponds roughly to the IPG frame 100 illustrated in Fig. 1. Of course, other layout configurations are contemplated to be within the scope of the present invention. For example, the program grid section 3301 may instead be on the right side, and the multimedia section 3302 may instead be on the left side. Similarly, the sections may
25 instead be on the top and bottom of an IPG frame.

In the embodiment depicted by the layout 3300 in Fig. 33, the program grid section 3301 comprises several horizontal stripes 3304-0 through 3304-7. The background shade (and/or color) may vary from stripe to stripe. For example, the background of some of the stripes may alternate from lighter to darker and so on.
30 Typically, the alternating backgrounds may be used to visually separate text information into channels or timeslots. For example, in the IPG frame 100 of Fig. 1, the alternating backgrounds of stripes 110-1 through 110-8 may be used to visually separate the program information into channels. Embodiments of the present invention may encode such

background stripes in such a way as to provide high viewing quality within a limited bit rate.

In accordance with an embodiment of the present invention, blank areas of the background are "skip" encoded to "save" a portion of the bit rate. In the example depicted in Fig. 33, the background for the program grid section 3301 which does not include any content other than constant color is skip encoded to save a portion of the bit rate for other uses.

Meanwhile, the quantizer stepsize for encoding the regions that include text is lowered to utilize the saved bits to improve the viewing quality of the text regions. The quantizer stepsize scales the granularity at which the image is quantized. Lower quantizer stepsize produces an increased fineness in granularity of the quantization. The increased fineness results in a higher viewing quality with lower loss of original content.

The quantizer step size chosen for each text region macroblock can be determined based on the rate allocated to the program grid portion. The program grid portion target rate is determined by subtracting the motion region 3302 target rate from the total bitrate. The program grid bitrate is then allocated to text and background regions by skip encoding the uniform color regions and then allocating the remaining bitrate to text regions via adjustment of the quantizer step size, e.g., MQANT parameter in MPEG-1/2. For text regions that show encoding artifacts, the quantizer step size is further forced to lower values.

In accordance with another embodiment of the present invention, the quantization matrix (also called the quantization weighting matrix) for encoding the program grid section may be optimized for encoding text, rather than being, for example, a standard or default quantization matrix. The MPEG compression standard, for example, provides two default quantization matrices: an intraframe quantization matrix for non-predicted blocks and an interframe quantization matrix for predicted blocks. The MPEG default matrix for non-predicted blocks is biased towards lower frequencies. The MPEG default matrix for predicted blocks is flat. A quantization matrix suitable for the specific program grid content is designed by analyzing the DCT coefficients of the transformed blocks. The coefficients are collected in a test pool and an optimum quantizer matrix is designed by a chosen rate-distortion optimization algorithm, which shall be known by a reader familiar in the art of quantizer design.

FIG. 34 depicts the program grid section 3301 of the layout 3300 of Fig. 33 in accordance with an embodiment of the present invention. As in Fig. 33, the

program grid section 3301 comprises several horizontal stripes 3304-0 through 3304-7. In the example depicted in Fig. 34, the stripes alternate from lighter to darker in order to visually delineate program information text into channels or timeslots.

5 In accordance with an embodiment of the present invention, encoding is performed on the program grid section such that encoded macroblocks do not cross a border between two stripes. In other words, stripe borders are aligned with the macroblocks in the program grid section. For example, as depicted in Fig. 34, each stripe 3304-X may be divided into three rows of macroblocks. The first stripe 3304-0 begins with a first indicated macroblock 3402, the second stripe 3304-1 begins with a first
10 indicated macroblock 3404, and so on. As shown in Fig. 34, the macroblocks do not cross any border between stripes. This avoids ringing and other defects that would otherwise occur if a macroblock crossed a lighter/darker border. The coding artifacts may appear at the border due to the high frequency edge structure of the stripe color transitions.

15 FIG. 35 depicts an encoding process 3500 that includes low-pass filtering in accordance with an embodiment of the present invention. The process 3500 is depicted in four steps.

The first step 3502 receives as input a source image and applies low-pass filtering. The low-pass filtering serves to reduce visual defects, such as ringing, because
20 those defects tend to comprise higher frequency components. The program guide grid high frequency components are removed, before the encoding process starts, to minimize the negative quantization effects of the encoder.

The second step 3504 receives the pre-filtered content and applies a forward transform to the source image. The forward transform may comprise, for
25 example, a discrete cosine transform. As a result of the forward transform, the image is transformed from image space to frequency space.

The third step 3506 receives the filtered output and applies quantization, as applied in MPEG-1/2 standards.

The fourth step 3508 receives the quantized output and applies lossless
30 encoding. The encoding may comprise, for example, a form of variable-length coding similar to the modified Huffman coding applied under the MPEG standard. An encoded image is output from this step 3508 for transmission to a decoder. In this method, the uniqueness of invention is the adjustment of the lowpass filter parameters in a certain

manner to remove the negative quantization effects of the quantizer in a pre-encoding stage.

The provided encoding optimization techniques can be applied within the context of slice-based encoding and picture-based encoding. Although various
5 embodiments that incorporate the teachings of the invention have been shown and described in detail herein, those skilled in the art can readily devise many other varied embodiments that still incorporate these teachings. For example, while some of the above figures depict horizontal background stripes, other embodiments of the invention may instead involve vertical background stripes. In addition, while a user interface with a
10 program information section is described above, other embodiments of the invention may involve other information sections. Similarly, while a multimedia section is described above, other embodiments of the invention may involve other display sections.

J. Music Interface for Media-Rich Interactive Program Guide

15 An aspect of the invention provides techniques for a music interface in a media-rich interactive program guide. The music interface can be implemented using a server-centric system, and can further be implemented as an extension to the user interface for programming guide for video channels.

FIG. 36 depicts a design of an interactive program guide (IPG) page 3600
20 in accordance with an embodiment of the invention. In the specific embodiment shown in FIG. 36, the IPG page 3600 includes a time slot region 3605, a guide region 3602, a program description region 3650, an icon region 3640, a video region 3620, a date/time display 3616, and a logo region 3615. Other designs for the IPG page with different layouts, configurations, and combinations of regions and objects can be contemplated and
25 are within the scope of the invention.

In the IPG page 3600, the guide region 3602 is used to display program listings for a group of channels. In the embodiment shown in FIG. 36, the program listings show the available programming in two half-hour timeslots. The program description region 3650 is used to present descriptive information relating to a particular
30 program selected from the program listings, or may be used to show other information. The video region 3620 may be used to display images, videos, text, or a combination thereof, which may be used for advertisements, previews, or other purposes. The video region 3620 may be implemented as described above in a server-centric manner. The logo region 3615 may include a logo of a service operator or other entity and may be

optionally displayed. The date/time display 3616 may be configurable by the user and may also be optionally displayed.

The icon region 3640 is used to display various icons, which may be created and/or enabled by the user. Each icon in the icon region 3640 can represent a filter or a link to another IPG page or a particular interface. Each filter selects a particular type of programming to be included in the program listings shown in the guide region 3602. For example, a Pay Per View (PPV) icon 3641 may be a filter that selects only PPV programming to be included in the program listings. A Favorites icon 3642 may be a filter that selects only channels designated by the user to be among his or her favorites. A Movies icon 3643 may be a filter that selects only movies or movie channels. A Kids icon 3644 may be a filter that selects only channels for children or programming appropriate for or produced for viewing by children. A Sports icon 3645 may be a filter that selects only sports channels or sports-related programming.

In an embodiment, a Music icon 3646 does not represent a filter. Rather, the Music icon 3646 is a link to a music interface, an embodiment of which is described in further detail below. An Options icon 3647 may also be a link to a menu of IPG options that the user may select amongst. The options may include (1) configuration and selection/deselection information of IPG related services, (2) custom information such as deactivating some of the filters or accessing the custom condensed listing menus, and others. A Weather icon 3648 may be a link to an interface to weather information.

FIG. 37 depicts a design of a music interface page 3700 in accordance with an embodiment of the invention. In the embodiment shown in FIG. 37, the music interface page 3700 includes a header region 3710, a left display region 3720a, a right display region 3720b, a channel description region 3750, and a logo region 3760. Again, other designs for the music interface page with different layouts, configurations, and combinations of regions and objects can be contemplated and are within the scope of the invention.

The header region 3710 is used to provide icons (objects), logos, text, graphics, and other information identifying the interface page being displayed. In the embodiment shown in FIG. 37, the header region 3710 for the music interface page 3700 includes a Guitar icon 3712, a Music logo 3714, a Saxophone icon 3716, and a Guide icon 3818.

In the embodiment shown in FIG. 37, each of the display regions 3720a and 3720b can be used to show music channel listings for a group of music channels.

These display regions 3720a and 3720b can be used to provide two columns of music channel listings. Each display region 3720 includes a number of rows, illustratively, 12 in FIG. 37, with each row corresponding to one music channel. Note that the "More" button in the display region 3720b provides a link to the next group of music channels, and the "More" button can be reached, for example, by the left, right, up, or down arrow keys on the remote control unit. Of course, the music channel listings may vary in format beyond that illustrated in FIG. 37.

The channel description region 3750 located at the bottom of the music interface page 3700 can be used to present descriptive information. In the embodiment shown in FIG. 37, the descriptive information relates to the music currently being played on a particular music channel (e.g., the channel selected by the user). The descriptive information may include, for example, the title of the song or music, the names of singers or musicians performing the song or music, the title of a compact disc or album that includes the song or music, the studio producing the music, and an identification number. Other or different descriptive information may also be presented. For example, a link (e.g., via an icon) may be included in the channel description region 3750 to provide an electronic commerce opportunity to purchase the compact disc or album or a digital (e.g., MP3) version of the song via download.

FIG. 38 depicts another design of a music interface page 3800 in accordance with another embodiment of the invention. The music interface page 3800 in FIG. 38 includes most of the regions shown in the music interface page 3700 in FIG. 37, except that the right display region 3820b is used to display a video or one or more still images or banners, instead of music channel listings. Similar to the channel description region 3850, the video or images displayed in the right display region 3820b may relate to the music currently being played on the selected music channel, or may be advertisements, previews of the music on a particular channel, or some other information. In another embodiment, the left display region 3820a can also be used to display a video or images instead of music channel listings. Generally, each of the display regions 3820a and 3820b can be used to display either music channel listings or video/images.

The contents of the music interface page can be sent via one transmission channel. Consider, for example, a 6 MHz transmission channel, with 27 Mbps of video, audio, data capacity. In an MPEG2 system, each audio channel is transmitted using 192 Kbps. Thus, for example, 44 music audio channels may be sent and occupy about 10

Mbps or less bandwidth, leaving around 17 Mbps of bandwidth to be possibly used for the video portion in the screen for images, videos, advertisements, barkers, and so on.

The music interface may be designed to function in various manners. In one embodiment, the user selects the Music icon 3646 on the IPG page 3600 and so enters the music interface page 3700 at a starting music channel, illustratively, channel 5 401. In another embodiment, the user may access the music interface page 3700 by directly entering a particular music channel number via the remote control unit or the STT. In this embodiment, the cursor is placed next to the selected music channel instead of the top of the music interface page. In yet another embodiment, the music interface page may be reached via a "spotlight" or channel information window. Thus, the music interface page can be reached via a number of different means. 10

In an embodiment, the user can move from one music channel row to another by pressing the up and down arrow keys or by directly entering the channel numbers. The right and left arrow keys in the remote control unit can transition the user 15 from one music channel to another in the same row (i.e., from the left display region to the right display region, and vice versa).

A particular music channel can be selected and tuned to in various manners. In one embodiment, as the cursor is moved to a music channel, the STT automatically tunes to the audio stream associated with that music channel. A data 20 stream associated with the audio stream may also be parsed and the data retrieved from the stream can be presented on the channel description region 3750 located at the bottom of the music interface page 3700, as depicted in FIG. 37. The retrieved data may include information about the audio channel content such as the title of the song and the CD information. In another embodiment, as the user scrolls through the music channel 25 listings, the current music channel is highlighted. The user can then select the music channel (e.g., by pressing a Select button on the remote control unit). In response, the STT tunes to the audio stream associated with the selected music channel. In yet another embodiment, the STT automatically tunes to a particular music channel if the cursor is left on that music channel for a particular period of time (e.g., 3 seconds).

30 As shown in FIG. 38, the left and right display regions 3820a and 3820b can each be used to display a video, one or more images, or banners. In an embodiment, after a particular time period (e.g., 10 seconds) in which a particular music channel in one of the display regions is selected and the STT has tuned to associated audio stream, the music channel listings in the opposite display region is replaced with a video or one or

more images related to the selected music channel. Alternatively, the opposite display region can be used to present advertisements, a banner, a preview clip, or some other information. The video to be displayed in the display region 3820 can be transmitted at a lower frame rate (e.g., six frames per second) and can be coded in various manners, as described in further detail below. The display region can revert back to showing music channel listings if the cursor is later moved into the display region.

The music interface page 3700 shown in FIG. 37 presents the music channel listings for one group of music channels. The listings for the next group of music channels can be selected for displayed by moving the cursor to the More button shown in the right display region 3720b or by pressing on a particular key (e.g., pressing the right arrow key twice) on the remote control unit. In response, another music interface page is generated and displayed. The user can move between the music interface pages by entering the More button, or a Previous button (not shown in FIG. 37). On each music interface page, the user can return to the IPG page 3600 by activating the Guide icon 3718 in the header region 3710.

The music interface pages 3700 and 3800 can be composed, rendered, transmitted, and generated in various manners. Referring to FIGS. 37 and 38, a number of icons and elements are used to generate the music interface pages. Some techniques to generate these icons and elements are described below.

For the embodiment shown in FIG. 37, the header region 3710 includes the Guitar icon 3712, the Music logo 3714, the Saxophone icon 3716, and the Guide icon 3718. Each of these icons and logo can be referred to as an object. In one embodiment, each object can be generated at the head end (e.g., as a bitmap) and pre-loaded to local memory of the STT, and can thereafter be provided for display whenever needed. In an embodiment, each object can be delivered to the STT as in-band data as part of private data in a video stream, after being composed as a bitmap and digitally encoded. Alternatively, the encoded bitmap data can be delivered via an out-of-band data stream that may be sent from the head end periodically or upon request by the STT. In another embodiment, the objects can be rendered at the head end and delivered to the STT via an out-of-band channel.

Each object in the header region 3710 can also be represented in an "enhanced" mode using a number of images (i.e., still pictures), a video, or some other representation. The video for the object (herein referred to as an "icon video") can be used to depict movement (e.g., rotation), animation, or some other effects. The icon

video can be used to "jazz up" the presentation of the object, and can thus enhance viewing experience. For example, the Guitar icon 3712 can change to depict different types of guitar, rotate about a vertical axis, and so on, and the Music logo 3714 can shimmer, change color, move up and down, and so on. The movement or animation of the objects can be activated (1) periodically at a regular time interval (e.g., every 20 seconds), (2) at all times, until turned off (e.g., by the user), (3) only when enabled (e.g., by the user, or upon the occurrence of some particular events such as a particular period of inactivity), or via some other means.

Each icon video can be rendered at the head end using slice-based encoding and sent via a particular PID. The decoder at the STT can retrieve and process the PID for each object to be animated. Alternatively, the images or video for each object can be composed as a set of bitmaps that are encoded and sent via an in-band or out-of-band channel. The STT can then process the bitmaps to generate the animation, which is then stored and later displayed when enabled.

To form the music interface pages shown in FIGS. 37 and 38, the STT may be designed and/or operated to generate a particular background for the page. The background for the music interface page can be designed to be similar to the background for the IPG page to give the same look and feel for the IPG and related interface pages. The background may include, for example, a number of horizontal colored stripes, with the strips alternating in color and each stripe corresponding to a music channel row.

The background can be generated in various manners. In one embodiment, the background strips are generated as a bitmap that is encoded and sent to the STT, which then receives, decodes, and draws the strips on the music interface page. The background strips can also be stored in a local memory of the STT. In another embodiment, the background strips are generated and sent continuously from the head end as in-band or out-of-band data.

In an embodiment, the music channel information (e.g., "401 Country Western," "402 Rock/Pop," and so on, in FIG. 37) may also be composed as a bitmap that is encoded and sent from the head end. In one embodiment, the bitmap data for the music channel information is delivered to the STT out-of-band via a separate stream or as part of a "spotlight" out-of-band data stream. In an alternative embodiment, the bitmap data can be sent in-band as part of private data in a video stream.

In yet another embodiment, the STT is provided with the channel map information and can render the channel information (e.g., "401 Country Western", and so on) to the proper location on the music interface page.

The descriptive information to be displayed in channel description region 3750 can be provided as an independent data stream and associated with the music channel information instead of the audio stream. This implementation may be used to present to the user a description for a particular music channel in the channel description region 3750 when the cursor is move to that channel row. The audio channel for that channel row can then be processed in response to the user hitting a Select key. In this manner, the user is able to first see the channel description (e.g., the title, CD, and so on) via cursor movement, and subsequently tune to the audio stream upon entering the Select key.

The video or images displayed in the display regions 3820a and 3820b can be rendered at the head end and provided to the STT. In an embodiment, the video is encoded as slices using the slice-based encoding techniques described above. Each video slice can then be sent as a separate stream (i.e., a separate PID). When directed, the decoder at the STT can retrieve and process the PIDs for the video slices and reassembles the slices to generate the video. The decoder can also generate the video at the desired location of the music interface page (e.g., the left display region 3820a or the right display region 3820b) by manipulating the slice-start codes of the video slices.

The PIDs for the video slices can be appropriately identified in a program map table related to the music channel number. For each music channel number, a distinct video clip, video, still images, banners can be associated and identified with a unique PID number. As the viewer selects a music channel, the decoder can retrieve the associated video PID.

In a similar manner, the channel rows in the music channel listings can be encoded as slices using the slice-based encoding techniques. Again, a number of PIDs can be used to send the music channel slices to the STT. The decoder can then generate the music interface page by retrieving the necessary PIDs and manipulating the slice-start codes of the music channel slices so that these slices can be appropriately reassembled (and possibly with the video slices) to generate the desired music interface page. Slice-based encoding, transmission, and recombination are described in further detail above and in the aforementioned U.S. Patent Application Serial No. 09/466,987.

In another embodiment, each of the display regions 3720a and 3720b in FIG. 37 is rendered at the head end as bitmap data as a complete region. As the user tunes to a music channel, one of the display regions become transparent and is replaced by the decoder with a related video.

5 In an alternative embodiment, each music interface page can be rendered as a video sequence at the head end and delivered via a digital video stream to the STT. The rendering and encoding of a music interface page as depicted in FIG. 37, at the head end, can be accomplished, for example, as described in the aforementioned U.S. Patent Application Serial Nos. 09/384,394 and 09/454,216.

10 The foregoing description of the preferred embodiments is provided to enable any person skilled in the art to make or use the present invention. Various modifications to these embodiments will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other embodiments without the use of the inventive faculty. Thus, the present invention is not intended to be limited
15 to the embodiments shown herein but is to be accorded the widest scope consistent with the principles and novel features disclosed herein.